AMSTERDAM STUDIES

IN THE

THEORY AND HISTORY

OF

LINGUISTIC SCIENCE

IV

CURRENT ISSUES IN LINGUISTIC THEORY

Volume 9

Current Issues in the Phonetic Sciences

Edited by Harry & Patricia Hollien

OFFPRINT

AMSTERDAM-JOHN BENJAMINS B.V.

1979

A STATISTICAL APPROACH TO THE PROBLEM OF ISOCHROMY IN SPOKEN BRITISH ENGLISH

D. R. HILL, W. JASSEM AND I. H. WITTEN
Calgary, Poznan and Colchester

It is difficult to reconcile actual measurements of durations in British English speech with text-book statements referring to a "tendency for rhythmic units based upon recurring word stresses to be of equal duration" -- so-called "theories of isochrony." There are a number of theories that attempt to explain these relationships; those of Abercrombie (1967), Halliday (1970) and Ladefoged (1975) for example -- all arise from the teachings of Jones (1918/1960). Moreover, an alternative rhythmic basis for isochrony had been formulated by Jassem (1952). For convenience, these theories will be referred to as the Halliday theory and the Jassem theory respectively.

The basis of the Halliday theory has been stated by Ladefoged (1975). He first introduces the notion of "stress" (or accent) and goes on to emphasize that a word having the potential for stress on a syllable when spoken in isolation may lose that stress in connected speech so that not all words have stressed syllables under these conditions. The rhythmic units in such a theory Halliday terms "feet," and a foot extends from one stress mark to the next. Marks inserted during "pauses," are termed "silent stresses," though it should be noted that we have, so far, found little evidence for their existence in our work. Ladefoged emphasizes that the stresses only **tend** to recur at regular intervals, but -- as a direct consequence of this -- it is necessarily asserted that feet tend to be of equal duration. Jassem's theory differs from Halliday's in that proclitic syllables (those belonging syntactically to the following foot) are excluded in computing the durations of his rhythmic units which he terms “rhythm units.” Thus, it should be noted that rhythm units and feet are identical if there are no proclitic syllables, otherwise rhythm units are shorter than the corresponding feet. The omitted syllables in Jassem1s terminology are dubbed “anacruses,“ by analogy with their counterparts in musical rhythm. Both approaches to describing the rhythm of spoken British English, however, postulate a tendency towards isochrony of the rhythmic units, and estimates of the degree of approximation to isochrony of rhythmic units might be expected to differ depending on whether feet or rhythm units were used in the analysis of a given body of data. Finally, it should be noted that rhythmic units fall into different classes. In a simple analysis, four classes may be considered: those containing a phrasal information point (tonic); those falling at the end of an utterance (final) ; those subject to both these conditions (tonic-final) and the remainder (unmarked).

METHOD

A tape-recorded sample of Educated British English speech -- i.e., Study Units 30 & 39 (Halliday (1970) -- was analysed spectrographically. The spectrograms were segmented by hand according to traditional phonetic criteria and the whole analysis (217 seconds) was subjected to double checking by a second reader. The segment identities and durations, together with certain other information (for example: tonic stress; syllable, word and foot boundaries; and syllable and foot types) were composed into computer readable form. Subsequently, the durations of syllables and both kinds of rhythmic unit (Halliday’s “feet” and Jassem’s “rhythm units”) were computed and statistically analysed. Simple models of syllable and rhythmic unit duration were also set up and run, the results being compared to actual measured durations. The qualified results of this analysis and modelling exercise, reported in our earlier paper (Hill, Witten & Jassem, 1978) may be summarized as follows: (1) Computer synthesis of spoken British English may be expected to produce reasonable approximations to the required rhythm.

(2) At least 25 percent of the determinants of rhythmic structure in spoken British English are unaccounted for in our simple model. (3) Despite a measured 6:1 ratio between the durations of the longest and shortest rhythmic units in our data, we have convincing evidence of a “tendency towards isochrony” that ranks third in importance as a determinant of segment duration assignment. Two questions now may be asked: (a) how can we quantify the notion of “tendency” in this context and (b) is there any reason to prefer one theory over the other.

DISCUSSION

One obvious quantity associated with any isochrony effect is the amount that it contributes to the total variance in mean segment duration, as in the part of this study that led to our formulation of a simple model for rhythm. Such an approach requires one to find out how much the variance in mean segment duration may be reduced by taking account of the size of rhythmic unit into which each segment falls -- assuming that decreasing segment durations are associated with increasing unit size. It was found that 9 percent of the variance could be accounted for by this factor, and that it was the third most important factor in determining segment durations in our model. It was also found that the equivalent syllable level effect did not exist -- we could find no evidence in our data of any syllable timing effect for spoken British English. However, two problems arise.

First, such a measure only tells us about that component of any tendency towards isochrony that arises from adjustment of segment durations -- what we may call “Segment Duration Compensation” or SDC. It is conceivable that spoken English could be strictly isochronous in the absence of SDC, if the speaker selected the semantics, syntax, words and phonetic realization of an utterance appropriately. Indeed, such selection is, perhaps, a large component of what a poet or lyricist does, although by no means always with the intention of introducing a tendency towards isochrony. Secondly, such a measure does not give a good “feel” for the amount of isochrony in utterances, even assuming that SDC were the only factor at work; and, furthermore, in our study, the measure does not distinguish at all between the two alternative formulations of isochrony in rhythmic structure.

An even more obvious approach would, perhaps, be to compute simple statistics on rhythmic unit durations. Table 1 shows the standard deviations for feet, rhythm units and anacruses as percentages of their respective means for SU30 and SU39. Again, it is difficult to gain a feeling for the amount of any tendency expressed by these figures, nor do they distinguish the two theories in any convincing fashion. However, the figures do represent all contributions toward equality of unit durations. We note here that anacruses show a great deal more variation than feet or rhythm units they appear to exhibit very little “tendency towards isochrony,“ yet they still exhibit a marked central tendency by this simple measure. It seems clear that almost any division of speech into units will appear more or less isochronous, if normalized standard deviation of unit size, or some similar statistic, is chosen as the measure.

A number of measures were investigated involving, typically, plots of average rate of segment production, or mean segment duration, against size of rhythmic units. Some were awkward to interpret because of the reciprocal relationships involved. The measure of the tendency towards isochrony finally chosen is based on the idealized plot found in Figure 1. In this plot, the relation of rhythmic unit duration to rhythmic unit size is compared for some class of rhythmic units. To ensure compatibility between plots of different sets of data, both axes are normalized -- the rhythmic unit duration being divided by mean rhythmic unit duration for the class, and the rhythmic unit size by mean rhythmic unit size. Thus the point (1,1) represents the mean of the data distribution, regardless of the data plotted, and the horizontal regression line must pass through this point (which implies that rhythmic unit duration is the same, regardless of size). A line through the origin, on the other hand, implies that the duration of

rhythmic units changes in strict proportion to the number of segments that it contains. Looking at this another way, a zero intercept on the y-axis implies no tendency towards isochrony based on SDC, while an intercept of 1.0 implies 100 percent isochrony, whatever the mechanism. A negative intercept would imply a disproportionate increase or decrease in rhythmic unit duration as its size changed, whilst an intercept exceeding 1.0 would imply that there was actually a decrease in the duration of rhythmic units as their size increased (i.e., overcompensation). This measure, although not entirely satisfactory, has many of the desirable properties of the required measure.

Our justification for choosing normalized plots as a basis for comparing different kinds of rhythmic unit, under different theories of “tendency towards isochrony,“ in terms of the degree of this tendency is simple and, hopefully, reasonable. The actual durations of rhythmic units under any theory are likely to be affected by a variety of factors, including speaking rate. If, as we found in various analyses of our original data, there are real differences in mean size and mean durations for different kinds of rhythmic units, even under the same theory, then some kind of normalization forms an essential precursor to comparison, or else such normalization will be implicit and hidden in any comparative analysis attempted. Theories based on the notion of isochrony merely assert a **tendency** for units to be more nearly equal than might be expected. If the average size and duration of different sets of units vary then it seems reasonable to express any measure of this tendency relative to the mean size and duration, just as one may describe the variation in scatter for different populations by referring to standard deviation as a proportion of the mean. In fact the method we have adopted in the proposed measure is very much of this character since it expresses the isochronous proportion of a set of rhythmic units as a percentage of the mean duration of the set.

Figures 2 to 5 show the plots obtained for each of the study units according to either theory. Similar type figures, but broken down by type of rhythmic unit, had to be omitted from this paper due to space limitations. Table 2 summarizes the degree of "Tendency Towards Isochrony (TTI) for SU3O and SU39 as: TTI = I x 100 percent -- where I is the y-intercept as described above.

Our figures show quite well something that was apparent in our modelling of British English rhythm -- namely that marked rhythmic units show a greater isochrony effect than unmarked units. However, they also show very clearly that the greatest effect is in final rhythmic units -- 60 to 69 percent of the mean duration in final rhythmic units is seen as fixed, regardless of size. At the other extreme, the anacruses (not, strictly, rhythmic units at all, but well worth analyzing in these terms) show a fixed component amounting to only 12 to 17 percent of their mean duration. Broadly speaking, anacruses show no significant SDC effect at all, and the correlation between length and size is high (e.g., 0.74 for SU30 as opposed to 0.46 for the final rhythmic units). However, there seems little other ground, at this level, for preferring the Halliday theory over the Jassem theory -- the former showing a slightly greater tendency towards isochrony on tonic feet, the latter on unmarked feet.

Further, it was noted from our plots that, although the slopes and intercepts for the two formulations are highly comparable, there is rather less scatter for the plots based on the Jassem formulation of rhythmic structure. This almost certainly relates to the exclusion of proclitic syllables, and may be the one reason, at the strictly acoustical level, for preferring the Jassem formulation of British English rhythm over the other. At this stage of the work, however, it has not proved possible to show that this, or any other difference between the two, is statistically significant. Nevertheless, the isochrony effect itself is significant, at better than the 1 percent level, whichever theory is chosen.

REFERENCES

Abercrombie, D. (1967). **Elements of General Phonetics**, Edinburgh. Edinburgh University Press, 96-97.

Halliday, M.A.K. (1970). **A Course in Spoken English: Intonation**, Oxford University Press, London, 1-3.

Hill, D. R. (1975). **Computer models for synthesising British English rhythm and intonation,** Proc 8th Int. Cong Phon.Sci

Hill, D. R., Witten, I. H. and Jassem, W. (1978). **Res. Rept 78/25/6**, Comp. Sci Dept., U. Calgary, Canada.

Jassem. W. (1952). Biul. Pol. Tow. Jezyk. 9:21-49.

Jones, D. (1960) **An Outline of English Phonetics**, Cambridge, Heffer, 237-242.

Ladefoged, P. (1975). **A Course in Phonetics**, New York, Harcourt, Brace, Jovanovich, 102-103.

ACKNOWLEDGEMENTS

The present study was supported by grant from A5261 from the National Research Council of Canada.

TABLES

Table 1: Standard deviations in rhythmic unit durations normalized as percentages of their respective means. (Figures in brackets denote mean durations in milliseconds.)

	SU30	SU39
Feet	33% (383)	37% (423)
Rhythm units	36% (311)	38% (356)
Anacruses	50% (138)	54% (181)

Table 2: Tendency towards isochrony (TTI)

	SU3O	SU39
All feet	54%	42%
All rhythm units	48%	44%
Final and tonic-final rhythmic units	69%	60%
Tonic feet	55%	50%
Tonic rhythm units	51%	44%
Unmarked feet	33%	28%
Unmarked rhythm units	37%	36%
Anacruses	17%	12%

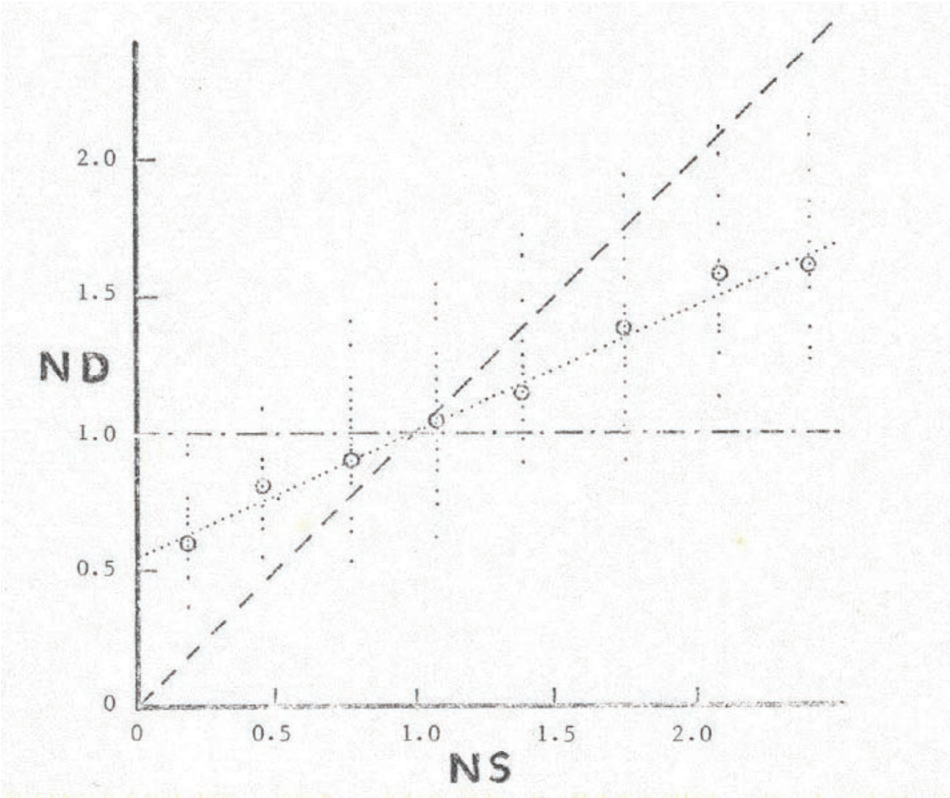

Figure 1: Generalised illustration of the relationship between normalised size (NS) and nomalised durations (ND) for hypothetical data

The horizontal chain-dotted line indicates constant rhythmic unit duration.
The dashed line at 45° indicates strict proportionality with size
The dotted line represents the actual regression line for the hypothetical data
Circles (o) indicate the mean of normalised durations for a given size (NS) category for the hypothetical data
(The dots scattered vertically suggest actual hypothetical data points)

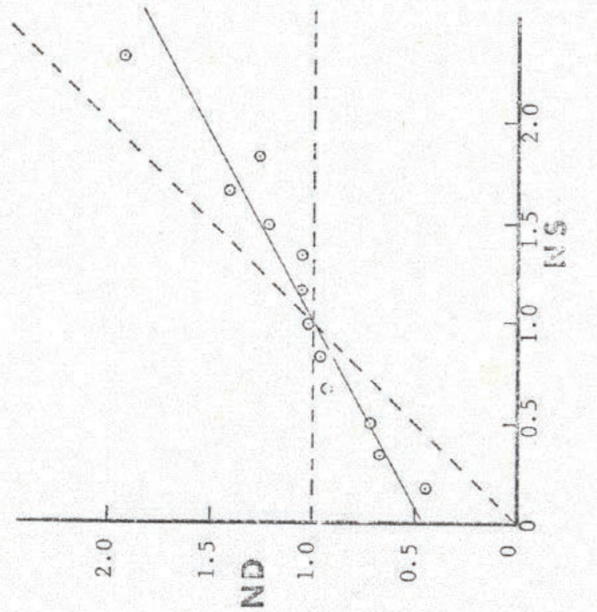

Figure 2: All feet for SU30

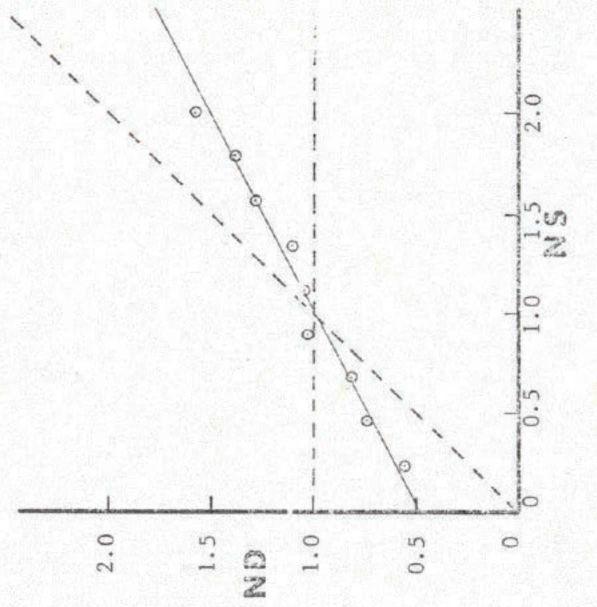

Figure 3: All rhythm units for SU30

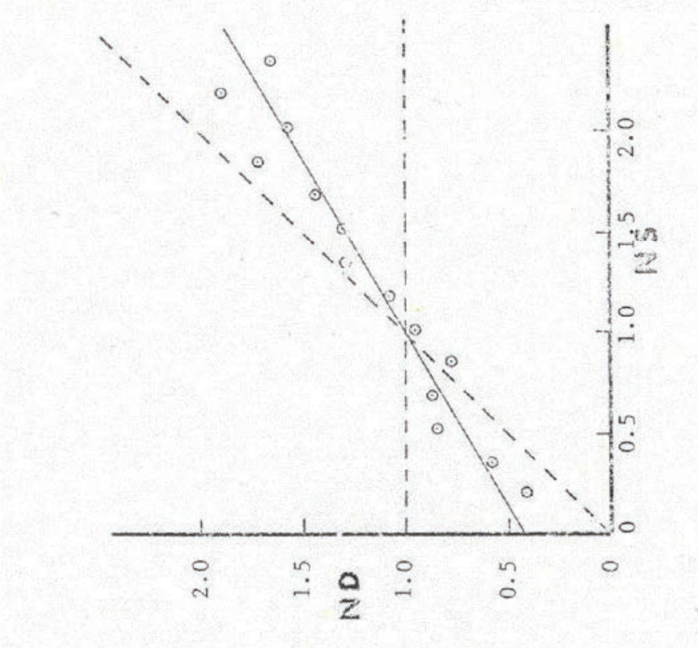

Figure 4: All feet for SU39

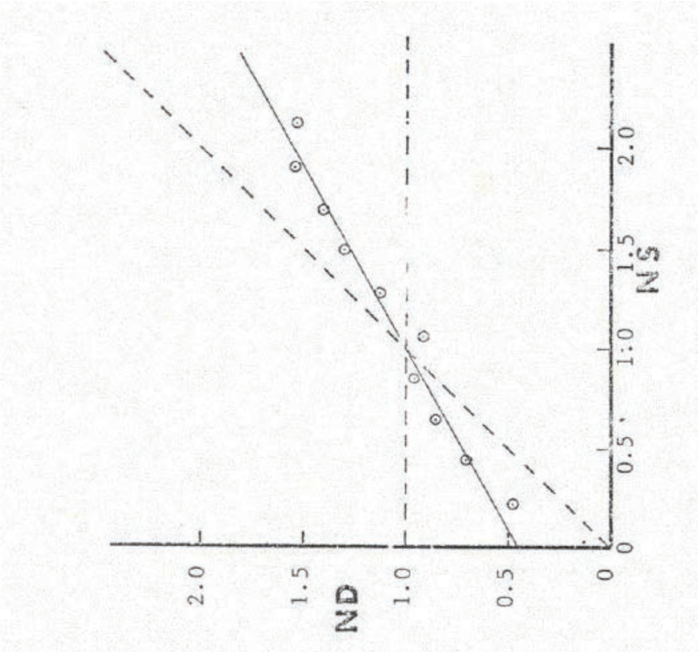

Figure 5: All rhythm units for SU39

CURRENT ISSUES IN LINGUISTIC THEORY

Since the spectrum of possibilities in linguistic theory construction is much broader and more variegated than students of linguistics have perhaps been led to believe, the *Current Issues in Linguistic Theory* (CILT) series has been established in order to provide a forum for the presentation and discussion of linguistic opinions of scholars who do not necessarily accept the prevailing mode of thought in linguistic science.

CILT is a theory-oriented series which welcomes contributions from scholars who have significant proposals to make towards the advancement of modern linguistic theory.

Current Issues in Linguistic Theory is especially designed, by offering an alternative outlet for meaningful contributions to the current linguistic debate, to furnish the community of linguists the diversity of opinion which a healthy discipline must have.